DOCUMENT RESUME

ED 084 682 88 EA 005 661

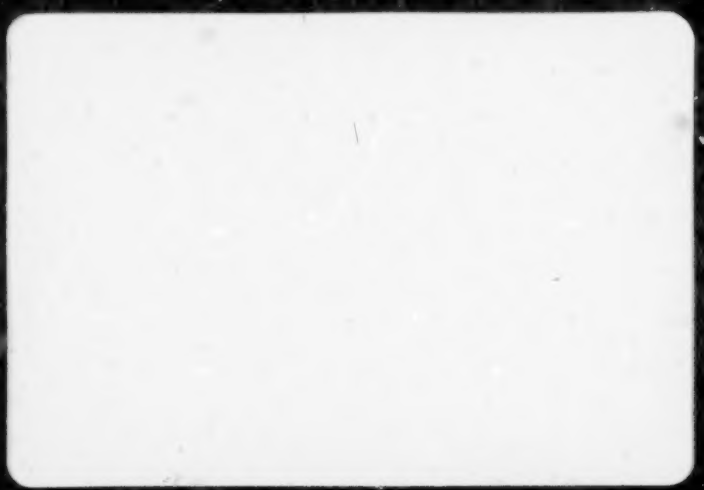
AUTHOR	Hall, Fred L.
TITLE	A Preliminary Evaluation of an Optimizing Technique for Use in Selecting New School Locations.
INSTITUTION	Chicago Board of Education, Ill. Dept. of Facilities Planning.
SPONS AGENCY	Bureau of Elementary and Secondary Education (DHEW/OE), Washington, D.C.
REPORT NO	71-7619
PUB DATE	Aug 73
GRANT	OEG-5-71-0078(290)
NOTE	37p.; Related documents are EA 005 330, EA 005 530, and EA 005 660
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Case Studies; Educational Planning; *Mathematical Models; *Planning (Facilities); *Programing; School Planning; *Simulation; *Site Selection; Student Distribution; Techniques; Urban Education
IDENTIFIERS	Elementary Secondary Education Act Title III; ESEA Title III; SIMU School

ABSTRACT

During the past two decades, mathematical programing techniques have been widely utilized in the private sector for optimization studies in locating industrial plants, scheduling commodity flows, determining product mix, etc. However, their use in the public sector has been less extensive, partly because of the absence of a clear-cut profit motive and partly because of the difficulty involved in expressing public planning problems strictly in terms of economic variables. This report describes a case study carried out in Chicago, in which an integer programing technique was used to investigate a basic problem in planning allocation of attendance areas. The study demonstrates that programing techniques can also be useful for planning public facilities such as schools. While the particular results reported are specific to the case study, the mathematical model described and tested is generic and can readily be adapted to other areas. (Author)

FILMED FROM BEST AVAILABLE COPY

ED 084682



ED 084682

This report was prepared pursuant to a grant from the U.S. Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A PRELIMINARY EVALUATION OF AN OPTIMIZING TECHNIQUE FOR USE IN SELECTING NEW SCHOOL LOCATIONS

By
Fred L. Hall
Assistant Professor
Department of Geography
McMaster University

Project Simu-School: Chicago Component

Funded by:
U.S. Office of Education
ESEA Title III, Section 306
Grant OEG-5-71-0078 (290)
Project 71-7619

Dr. Joseph P. Hannon
Project Director

Simu-School: Center for Urban Educational Planning
Chicago Board of Education
28 East Huron Street
Chicago, Illinois 60611

August, 1973

FOREWORD

Modern-day educational planners face an extremely difficult task of providing quality education to large masses of students in view of decreased revenues, soaring costs, shifting populations and changing educational programs. Such a challenge requires that a far greater emphasis be placed on planning for schools than has been the case to date and necessitates the development of improved techniques specially designed for educational planning.

Project Simu-School is intended to provide an action-oriented organizational and functional framework necessary for tackling the problems of modern-day educational planning. It was conceived by a task force of the National Committee on Architecture for Education of the American Institute of Architects, working in conjunction with the Council of Educational Facility Planners. The national project is comprised of a network of component centers located in different parts of the country.

The main objective of the Chicago component is to develop a Center for Urban Educational Planning designed to bring a variety of people--laymen as well as experts--together in a joint effort to plan for new forms of education in their communities. The Center is intended to serve several different functions including research and development, investigation of alternative strategies in actual planning problems, community involvement, and dissemination of project reports.

During the past two decades, mathematical programming techniques have been widely utilized in the private sector for optimization studies in locating industrial plants, scheduling commodity flows, determining product mix, etc. However, their use in the public sector has not been quite as extensive, partly because of the absence of a clear-cut profit motive and partly because of the difficulty involved in expressing public planning problems strictly in terms of economic variables. This report describes a case study carried out in Chicago in which an integer programming technique was used to investigate a basic problem in planning educational facilities: the optimal location of schools and the optimal allocation of attendance areas. The study demonstrates that programming techniques can also be useful for planning public facilities such as schools. While the particular results reported on the following pages are specific to the case study, the mathematical model described and tested is generic and can readily be adapted to other areas. It is hoped, therefore, that educational facility planners will find this report interesting and useful.

Ashraf S. Manji
Project Manager

CONTENTS

	Page
Introduction	1
Mathematical programming	3
A case study	8
Potential planning objectives	8
Selection of the case study	10
Analytical procedures	14
A representative solution	16
Discussion of findings	20
Summary and implications	27
Evaluation of the model	29
Implications for policy	31
Notes	33

A PRELIMINARY EVALUATION OF AN OPTIMIZING TECHNIQUE FOR USE IN SELECTING NEW SCHOOL LOCATIONS

INTRODUCTION

Determining locations for new schools is a problem which has not received much analytical attention. Studies have been made of the criteria involved in the selection of school sites, but location factors have played a very minor role in these.[1] There have been a number of statements about how schools should be located in relation to the students they will serve, but these are for the most part quite general, and often somewhat contradictory. For example:

> Schools . . . should be located near the center of the present and probable future school population. It is desirable, whenever it is possible, to locate schools within walking distance of the greatest number of pupils . . . .
>
> School boards should not lose sight of the fact that transportation to and from school over a long period of years is a significant cost item. Locating a school on a site requiring pupils to travel long distances is questionable economy of time and money and should be avoided where feasible.[2]

Several possible location criteria are listed here--travel time, monetary costs of travel, and "within walking distance of the greatest number." Are these all consistent with each other? If each would lead to a different choice of location, which is most important? Furthermore, the same criteria which apply to the selection of school locations should also apply to the determination of school attendance areas. The criteria have not been analyzed thoroughly in that context either.

This paper reports on an investigation of a number of potential criteria for this combined location-allocation problem (i.e. the problem of locating new schools and of allocating pupils to all schools). At the same time, it reports on an analytical model which was developed to carry out this investigation, and which appears to have considerable potential for general use in school planning studies.

Because the model is basic to the study, and represents an addition to the set of tools available to the educational facility planner, it is described first in the following report. After discussion of the model in general terms, the specific case study is presented, including details of the criteria which were investigated, a description of the case study area, and findings which were derived from the case study about the location criteria. Conclusions about the model are presented in the final section of the report, along with an overview of the study and its implications.

MATHEMATICAL PROGRAMMING

In general, mathematical programming represents one approach to the solution of constrained optimization problems. That is, it deals with problems in which one is attempting to optimize (either maximize or minimize) some explicit objective, subject to a number of limitations or constraints on combinations of the variables involved. The earliest advances in programming were made with regard to problems in which the objectives and constraints could be expressed as linear functions of the variables. This area of programming, called linear programming, has been used since as early as 1963 to help delineate school districts.[3]

A typical districting problem can be stated as follows. Assign students to schools in such a way as to minimize the total amount of travel necessary for all students to get to schools, subject to these two constraints: (1) every student must be assigned to one and only one school; and (2) no school can be assigned more students than it has capacity. The decision variable in this problem is the assignment of a student to a school, or more often, the assignment of a stated proportion from one census tract to a particular school. This problem can be stated mathematically very simply. Find the set of x_{ij} which will

$$\text{minimize} \qquad Z = \sum_i \sum_j d_{ij} x_{ij}$$

$$\text{subject to} \quad (1) \quad \sum_j x_{ij} = 1 \qquad \text{for all tracts, } i$$

$$(2) \quad \sum_i p_i x_{ij} \leq c_j \qquad \text{for all schools, } j$$

$$x_{ij} \geq 0 \qquad \text{for all combinations of tracts, } i, \text{ and schools, } j$$

where x_{ij} = the fraction of students from tract i attending school j

d_{ij} = the distance from tract i to school j

p_i = the student population of tract i, and

c_j = the capacity of school j.

The first three lines represent the objective function and two constraints listed above; the last line ensures that all assignments will be positive. Note that these are linear relations: the variables in all the equations are not raised to any power nor are there any products of two or more x_{ij}. Solution procedures for linear programs are quite well developed and have been for a number of years, so that many applications of them have been solved. For example, this basic linear program, or a minor variation of it, has formed the basis for a number of recent papers dealing with re-districting for racial balance.[4]

In the simplest solutions to the basic problem, most of the variables, x_{ij}, will be equal to zero. For a given tract, i, either only a single x_{ij} will be non-zero, in which case that x_{ij} will equal 1.0 and all students from the tract will attend the same school, or perhaps two or three will be non-zero, and the pupils from that tract will be assigned to several schools. In the applications to re-districting for racial balance, more tracts usually receive split assignments, but still only to two or three schools.

Unfortunately, linear programming is of use only for determining districts for existing schools. It does not help when one is attempting to locate new schools, because of the nature of the decision variables involved. They deal with assignment of students to schools, but if it is uncertain where the schools are, then such variables are of little use. A different type of decision variable is needed.

Recent computational advances in another aspect of mathematical programming provide the opportunity to use an additional type of decision variable. This is the field of integer programming, in which the variables can take on only integer values, namely the values zero (0) or one (1). Using one such variable for each potential new school location, the variable will take the value one (1)

if a school is to be built there, and zero (0) if a school is not to be built in that location. With this decision variable to determine where the schools are located, it is possible, within the same programming framework, to use the x_{ij} variables to determine school assignments as before. The result is similar to a linear program, as described previously, but includes a number of integer variables as well.

The verbal formulation of this location-allocation problem for new schools can be given as follows. Determine locations for new schools, and the resulting allocation of students to all schools, both old and new, so that the result minimizes the total distance travelled to school by all the students, subject to the following constraints:

(1) each student must be assigned to one and only one school;

(2) no school can be assigned more students than its capacity, and each school, new or old, must be assigned those students living in the tract in which it is located;

(3) a specified number of new schools is to be built.

The addition of the restrictive assignment under the second constraint is not necessary for the program, although it does simplify it somewhat. It was introduced primarily to ensure that students living next to a school would not be assigned to some more distant school. As will be explained later, numerous modifications of this type are possible within the basic programming framework. For example, it would be possible to extend such a restriction so that all students living within one-half mile of a school must attend that particular school.

The mathematical formulation of the location-allocation problem is quite similar to the formulation of the simple linear programming allocation problem. The main difference is that the self-assignment variables, x_{jj}, of the old formulation are now used to represent the integer variables determining the new

school locations, as well as to indicate the fact of assignment to a school within the tract. A second difference is the introduction of the third constraint, limiting the number of new schools. The problem is now to determine those values of the x_{ij} (including the x_{jj}) which will

$$\text{minimize} \qquad Z = \sum_i \sum_j d_{ij}\, x_{ij}$$

$$\text{subject to} \qquad (1) \quad \sum_j x_{ij} = 1 \qquad \text{for all tracts, } i$$

$$(2a) \quad \sum_{\substack{i \\ i \neq j}} p_i\, x_{ij} \leq (c_j - p_j) + c x_{jj} \qquad \text{for those tracts, } j, \text{ which contain an existing school}$$

$$(2b) \quad \sum_{\substack{i \\ i \neq j}} p_i\, x_{ij} \leq (c - p_j)\, x_{jj} \qquad \text{for all tracts, } j, \text{ without an existing school}$$

$$(3) \quad \sum_j x_{jj} = m$$

$$x_{ij} \geq 0$$

$$x_{jj} = 0,1$$

where, as beofre,

x_{ij} = the fraction of students from tract i attending school j

d_{ij} = the distance between tract i and school j

p_i = the student population of tract i, and

c_j = the capacity of existing school j

and the new symbols are

x_{jj} = the integer decision variable for new school locations

c = the capacity of a new school

m = the number of new schools to be built.

It is necessary to treat differently those tracts with and without existing schools. This is because we stated that all tracts with schools, new or old, must self-assign. For tracts with existing schools, the tracts' population must be subtracted from existing capacity (equation 2a); for those without schools at present, the population will be subtracted from potential new capacity. As formulated, the problem allows new capacity to be added to existing schools. If these schools are well-placed, it may indeed be best to build an addition, rather than to build a separate school at a new location. Certainly the problem formulation should permit such a possibility, so that it can be adequately tested.

To assist in interpreting this formulation, consider two tracts, neither of which has an existing school. Assume that, in the solution, tract 1 does not obtain a new school, but tract 2 does. Then x_{11} will be equal to zero, and x_{22} will be equal to unity. Consequently, equations (1) ensure that x_{2j} will be zero, for all schools, j, other than that in tract 2; while it is still necessary to find some positive values of x_{1j} to sum to unity. Equations (2b) ensure that, for tract 1, no assignment to it is possible, since the right hand side of the inequality is equal to zero; and that for tract 2, there is positive capacity of $c - p_2$ to be filled. Hence it is possible for x_{12} to equal unity, and all of tract 1 to be assigned to the new school in tract 2.

A CASE STUDY

Two purposes governed the application of the mixed integer programming location model to a case study. The original reason for constructing the model was to permit an examination of some of the different criteria which have been suggested for selecting high school locations, and this interest continued to be primary in the case study. Secondarily, the case study was also intended to provide information on the usefulness of this type of model for location questions. This section of the report begins with a discussion of the several location criteria which have been suggested for the school problem. Following a description of the case study, and the rationale for its selection, it then goes on to present the findings from the case study, including an example of the kind of output provided by the model.

Potential planning objectives

Although all of the previous applications of programming methods to school problems have been concerned with minimizing the total travel by students to and from school, there has been little agreement on how best to measure travel. At least four different measures have been used: distance; time; monetary cost; and the percentage of students who must take a bus to school. While all four measures have been in use for some time, very few authors provide any explanation for the selection of one rather than another. Indeed, few even appear to consider the possibility of using different measures. One purpose of this case study was to apply these different travel measures to the same problem, to see if there is a best one to use in the school planning context. The principal questions investigated dealt with the implications of each measure for the resulting spatial pattern, and with the behavior of the other three travel measures when a particular one was minimized.

In addition to considerations of total travel, suggestions have occasionally been made that there should be limits placed on the amount of travel any individual is required to make to attend school. In fact some studies have proposed attempting to minimize the maximum travel necessary for any individual in the system. In the case study, these considerations were applied as constraints on the maximum travel by any student (e.g. no student may travel more than three miles to school), and the trade-offs between total travel costs for all students and these individual limits were examined, as were the spatial implications of these constraints.

Racial desegregation has also been of considerable concern with regard to school planning and districting in the past several years. However, little has been done to determine the increased travel costs brought about by increased desegregation. The programming model provided a useful procedure to investigate these trade-offs between increased levels of desegregation and increases in travel costs. It was hoped that such an analysis might uncover some level of desegregation at which there was a sudden steep increase in travel costs, so that strong economic arguments could be made for achieving that particular level of desegregation.

Extensive searches of a number of literature areas--including those on school planning, general public facility planning, and traditional private sector location theory--indicated that these two kinds of objectives, travel and desegregation, were the only ones directly applicable to the school location problem. A variety of other concerns were mentioned relating to school site selection; but these were not really location attributes in the sense that the term is being used here (i.e. location with respect to a population being served). The aim of the case study was to investigate these two kinds of objectives, and to determine how useful the mixed integer programming approach is for obtaining information about them.

Selection of the case study

Two issues arise in selecting a case study. First, an appropriate area must be found for testing the model, preferably an area in which the issues addressed by the model are real and present problems, and one which meets several qualifications of size and representativeness to be explained below. Once an appropriate area has been found, the question of the duration of the planning horizon must be settled. On both of these issues, our final choice proved more restrictive than we had hoped, but the reasons for limiting the study were extremely persuasive.

The area selected for study was one of the administrative districts in the city of Chicago-- District 18, in the southwest part of the city. On the one hand, the number of public high school students in this one district in Chicago (7,874 in 1970) was greater than the public high school enrollment in all but one other city in the state of Illinois (Rockford, with 11,891). The results from this study area should therefore demonstrate the applicability of the model to cities of quite reasonable size. On the other hand, District 18 was small enough that it held promise of keeping the cost of solution low enough to allow a dozen or more variations to be solved, permitting the investigation of the location criteria which was the main purpose of the study. Additional factors in favor of District 18 included the fact that the racial groups in the district are residentially segregated, which is typical of most cities, and the fact that the district does not have a compact shape, but is instead rather irregularly shaped (Exhibit 1).

Selection of a realistic planning horizon presented more formidable problems. Ultimately, the decision was made to use the model on present data only, despite the obvious lack of realism in such a case study. However, in order to use other than present data, three problems would have to be

solved: the matter of an appropriate time horizon; the problem of population prediction for small areas; and the question of how to identify optimality over a span of years. As each of these issues presents a major problem by itself, it was decided to use available, present data. If the model proved effective with these data, it would be equally effective with any similar set of data for future years. Hence, although this decision limits the realism of the actual output from the case study, it does not represent a limitation of the model. Rather, it shows the limitations of time and funds available for this particular application.

At the time this study was begun, only preliminary results were available from the 1970 census. Using these, plus information on 1970 high school enrollments, and information on parochial school enrollments in the vicinity, we arrived at the distribution of public high school students by race shown in Exhibit 1.[5] As is apparent in the diagram, de facto residential segregation is present in the district although there is no single strong racial boundary. The public high school enrollment in the district is about 55% Negro, considerably higher than normal, but it was felt that this was unlikely to affect the generality of the results.

The three existing high schools are located as shown on Exhibit 1: Carver, with a nominal capacity of 800, in tract 5401; Fenger, capacity 2100, in tract 4912; and Morgan Park, capacity 2170, in tract 7502. In 1970, the enrollments of these schools were, respectively, 1082, 3155, and 2355. An additional 1282 students attended branches of Fenger and Morgan Park set up in various elementary schools. The total nominal capacity of the three schools is 5070; the total high school enrollment in 1970 was 7874. Obviously, new school capacity was needed in the district: the location problem discussed here is not solely of academic interest. It was decided to attempt to locate two new facilities in

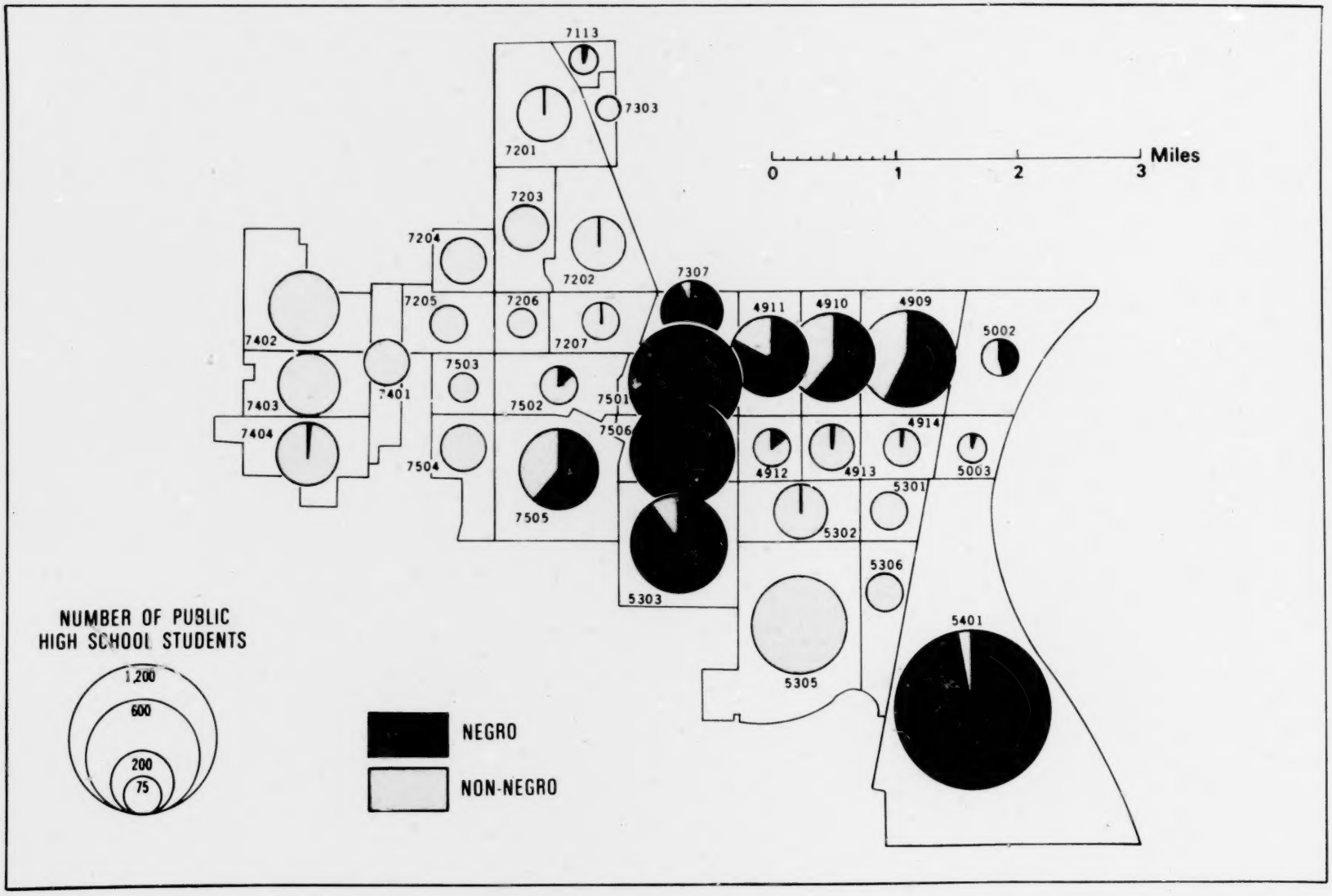


Exhibit 1. Distribution of public high school students in the study area, by census tract.

the district, each to serve 1500 students. (The schools could of course be built for more students, but we were locating them in terms of the present population, and desired each to serve no more than 1500 from that group.)

The structure of the model led us to treat each census tract as a point source of population, thereby assuming that all of the population in a particular tract originated from that point. In addition, if the tract contained a school, the school was assumed to be located at that central point as well. These assumptions very much simplified measuring the transportation costs, and helped to keep the model to a reasonable size.

The location criterion relating to racial integration can be handled on the basis of the information so far described. The transportation related criteria require additional data--on distances, travel times, and travel costs within the study area. Distances between the population points were calculated on a North-South and East-West rectangular street grid, because this best represents the pattern of streets in the district. The travel time and monetary cost measures were based on the type of publicly provided transportation available in the district. In the study area, as in most of Chicago, students ride regularly scheduled Chicago Transit Authority (CTA) buses if the distance to school is too great to walk. This usage of CTA buses is subsidized, and no other publicly provided transportation is available.

It was assumed that students can walk no more than 1 1/2 miles to school--either directly, or from home to a bus line and from that to the school. Walking speed was taken to be three miles per hour, bus speed ten miles per hour, and one transfer between buses was permitted. Average waiting times for the buses used were also included. A computer program was written to calculate the minimum time path between all pairs of census points, and the output from this was used as the travel time data for the analysis.

The measure of monetary travel cost selected was the out-of-pocket cost to the student. The primary reason for this choice was that the total dollar output was comprised of CTA costs, state subsidies, and fares paid, and that to sort all that out was too complicated for a first-pass examination of the model or the location criteria. As a result, there were only three levels of cost: zero if the student walked; twenty cents if he rode one bus; and thirty cents if he transferred to a second bus.

The measure of the percentage of students bused was calculated manually after solutions were obtained, because it was discovered too late to be entered into any of the computer runs.

Analytical procedures

Using the data detailed above, eleven variations of the basic model were solved. As discussed in the section on mathematical programming the model was structured in such a way that measures of total transportation entered as the function to be minimized. The criterion relating to maximum individual travel appeared essentially as a constraint, but its actual effect was to reduce the size of the problem by removing from consideration all potential origin-destination pairs for which the travel was above the set limit. The integration criterion appeared as additional constraints on the solution, specifying that the number of black students at a school could not be more than a certain limit, nor less than another limit. For tracts without existing schools, the constraints were of the form

$$\sum_{\substack{i \\ i \neq j}} b_i \, p_i \, x_{ij} \leq [(B + v) \, c - b_j \, p_j] \, x_{jj} \qquad \text{and}$$

$$\sum_{\substack{i \\ i \neq j}} b_i \, p_i \, x_{ij} \geq [(B - v) \, c - b_j \, p_j] \, x_{jj}$$

where, as before,

x_{ij} = proportion of students from i attending school at j,

p_i = student population of tract i, and

c = capacity of the new school.

The new terms necessary for the racial criterion are

b_i = the fraction of the student population of tract i which is black,

B = the fraction of the district student population which is black, and

v = the amount by which the racial mix at each school is allowed to vary from the district average.

If, for example, v is set at 15%, these two constraints ensure that the racial mix at each school will be between 70% and 40% black (i.e. 55%, the district ratio, plus or minus 15%).

Seven of the variations of the model focused on the degree of racial integration achieved. In these, time was used to measure total travel, no limits were placed on individual travel, and two new schools were added, each with a capacity of 1500 students. The variable v (in the racial integration constraints just described) took on a different value in each variation--one, five, ten, fifteen, twenty, and twenty-five percent; plus one run in which the racial constraint was totally ignored, equivalent to a value of v equal to fifty-five percent. Total travel on the remaining three measures (distance, dollars, and percentage bused) was calculated for each of the solutions to these variations, although only the time measure was actually minimized.

Two variations investigated the differences in travel costs and spatial patterns produced when different travel measures--distance and dollars--were minimized. For these runs, the permitted racial variation was kept at fifteen percent, two schools (capacity 1500) were added, and no individual travel limits were in effect. This enabled a comparison with the earlier run in which v was set at fifteen percent, but time was minimized.

The final two runs dealt with implications of upper limits on individual travel for optimal spatial patterns and for total travel costs. A maximum distance of three miles was used for one run, and a maximum travel time of thirty minutes was employed in the other. In both runs, total travel time was the function to be minimized; the permitted racial variation was fifteen percent; and two new schools were added. This choice of parameter set allowed direct comparison of these results with the case in which time was minimized, but no individual travel limits were in effect.

A representative solution

To provide some indication of the kinds of output produced by the model, the solution for one of the runs is presented here. Because it appears in the analysis of all three location criteria, the most important run to examine is the one which minimizes total travel time with no maximum travel constraints and allows a fifteen percent variation in the racial mix. That is, we shall discuss in this section the optimal solution to the following problem: determine optimal locations for two new schools (each of capacity 1500), and the resulting allocation of students to all schools, in such a way that the total travel to school by all students is minimized, subject to the constraint that each school must have a Negro enrollment which is between forty and seventy percent of the total enrollment.

Exhibit 2 shows the values of all the non-zero variables for the optimal solution to this problem. The first three columns contain assignments to existing schools; the last two columns denote the locations of the new schools and the allocations of students to them. The first five rows of the table are easily interpreted: all of the students from each tract go to a single school. Those from tracts 4909 and 4912 go to the school (new and old, respectively)

Exhibit 2

Representative optimal solution

Problem description:

objective function = time
permitted racial variation = 15%
new schools = 2
no travel constraints

From \ To	Fenger	Carver	Morgan Park	5303	4909
4909	.	.	.	.	new
4910	.	.	.	.	100
4911	100	.	.	.	.
4912	old	.	.	.	.
4913	100	.	.	.	.
4914	.	6	.	.	94
5002	.	.	.	.	100
5003	7	93	.	.	.
5301	.	100	.	.	.
5302	99	1	.	.	.
5303	.	.	.	new	.
5305	26	.	.	74	.
5306	.	100	.	.	.
5401	11	.	.	.	89
7113	100	.	.	.	.
7201	100	.	.	.	.
7202	100	.	.	.	.
7203	.	.	100	.	.
7204	.	.	100	.	.
7205	100	.	.	.	.
7206	.	.	100	.	.
7207	.	.	100	.	.
7303	.	.	100	.	.
7307	72	.	28	.	.
7401	.	.	100	.	.
7402	45	.	55	.	.
7403	.	.	100	.	.
7404	.	.	100	.	.
7501	.	.	100	.	.
7502	.	.	old	.	.
7503	.	.	100	.	.
7504	.	.	.	100	.
7505	.	.	100	.	.
7506	100	.	.	.	.

Total travel measures

time (student-minutes)	137,099	
distance (student-miles)	10,439	
dollars	709.0	
percentage bused	37.64	
Desegregation index	94.40	% of most complete possible

within that tract, by definition. Tract 4910 is assigned to the new school in tract 4909, and tracts 4911 and 4913 to the existing Fenger H. S. The sixth row shows that a split assignment is necessary for tract 4914: six percent of its students will attend Carver H.S., and ninety-four percent will attend the new school in tract 4909. The remainder of the table is read similarly.

The bottom portion of the table shows the calculations for the measures of total travel. Only the time measure represents the best possible value. That is, we can say with certainty that it is impossible to get all students to school in less than a total of 137,099 minutes, if we wish to maintain a fifteen percent racial variation, and to add two new schools. The other measures listed have not been optimized: they simply report what the best time solution implies for the other measures and criterion. Several of these numbers can be interpreted more easily if placed on a per-student basis. The average travel time is roughly 17 1/2 minutes per student, each way; average distance traveled is about 1 1/3 miles; and average out-of-pocket cost is 9¢ per student.

The allocation results listed in Exhibit 2 have been mapped in Exhibit 3. It is readily apparent from this figure that minimal travel times do not give rise to compact, contiguous attendance areas when one attempts to desegregate a district such as this. Consider the Fenger attendance area, for example. There is a contiguous area near the school assigned there, but there are also four other isolated portions of its attendance area, scattered throughout almost all parts of the district. While that is the extreme case in this solution, none of the other attendance areas are totally contiguous, either. However, questions of contiguity, or of long individual trips, were not explicitly considered in this formulation of the problem. If they are felt to be important, they can be introduced.

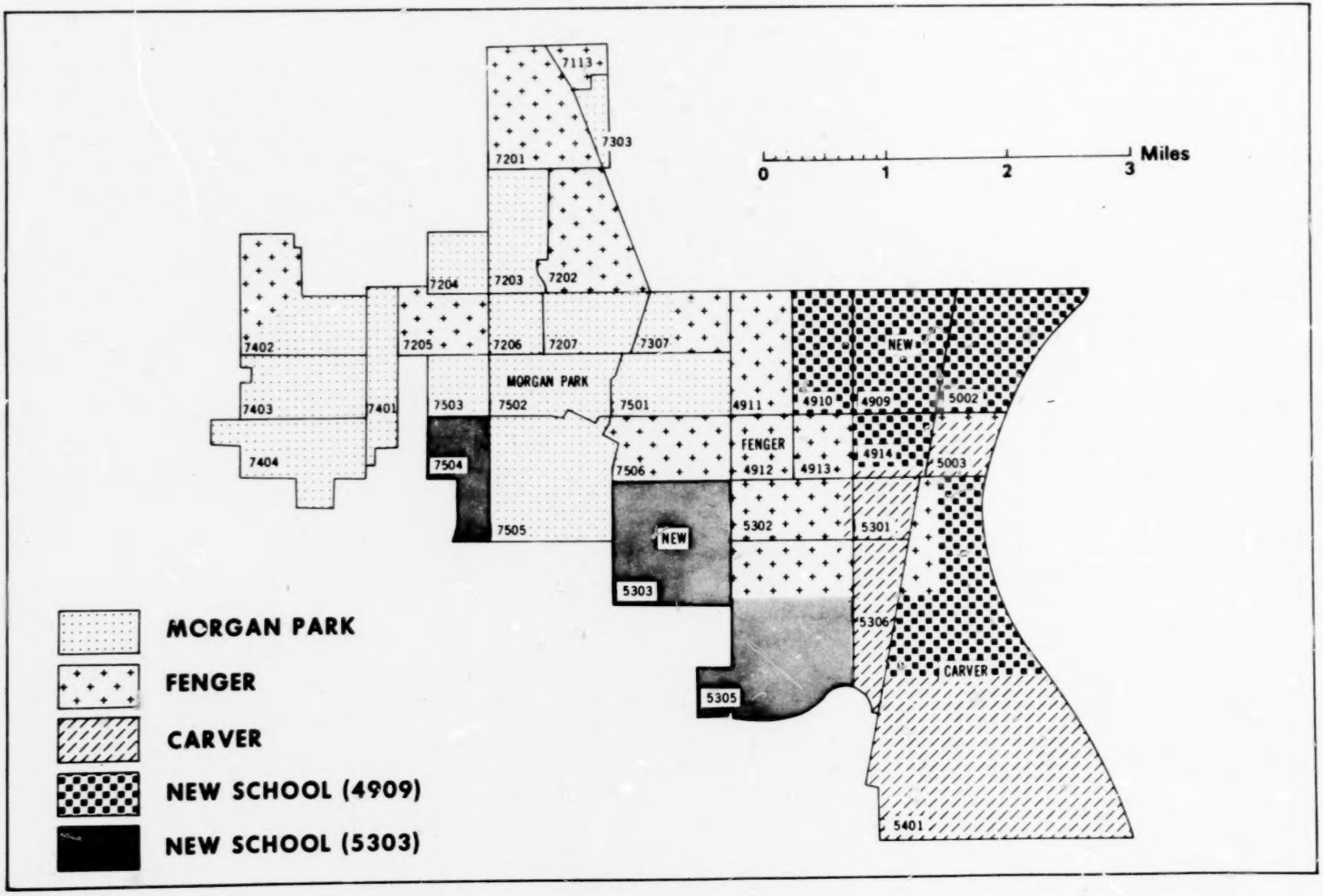


Exhibit 3. Optimal solution when travel time is the objective function.

Discussion of findings

The case study was selected in order to obtain information about location criteria and about the model. This portion of the report discusses the findings related to the location criteria. Although specific statements and numbers will apply only to this particular study area, the general conclusions have a wider validity as well. In addition, the discussion here gives some idea of the range of findings possible with the model, and of the ways it might be applied in the future. Specific evaluation of the model, however, is reserved for the final section of this report. The present section focuses on three aspects of the location criteria: the four measures of total transportation costs; the two measures of integration and the trade-offs between them and transportation; and the relation between total travel and limits on individual travel.

Measures of total travel

Analysis of the measures pertaining to total travel was originally intended to identify a single best measure for use in school location decisions. To accomplish this, two modes of investigation were used: inspection of the spatial patterns produced by each measure; and comparisons of the measures over all solutions, to identify functional similarities among them. The findings indicate that the choice of an optimization measure makes a significant difference to the resulting location pattern, but that no single measure stands out as best.

The investigation of spatial patterns showed that, for all three measures used, both new schools were located in the eastern half of the district, (as shown, for example, in Exhibit 3). In view of the distribution of the public high school population in the district (Exhibit 1), this similarity

of locations is not particularly striking. No matter what cost measure is used, it seems reasonable to expect new locations to be chosen within the more densely populated area. In fact, given this distribution of population, differences in the location choices become more important. Not only do locations differ under alternative travel measures--the allocation patterns, of students to schools, show considerable variation from one measure to the next. There is no way to say which pattern is best, hence no way to identify a best measure for planning, although it is obvious that the measures produce differing results.

Comparison of the actual measures calculated for each solution further supports this finding. When time was minimized, the resulting solution entailed costs of 137,099 minutes, 10,439 miles, and $709. The minimum distance solution entailed an increase of 13.6% in total time, to 155,708 minutes, while decreasing the travel distance to 9,910 miles. The monetary cost of this solution was $643.40. The solution which minimized monetary costs entailed one-way daily out-of-pocket costs of only $455.60. (The distance-minimizing solution represented an increase of 41.2% over this; the time-minimizing solution, a 55.6% increase.) However, total distance for this solution increased 16.4% over the lowest possible, to 11,536 miles; and total time increased 22.5% from the minimum to 168,068 minutes. Thus minimizing one measure tends to increase the others, and again selection of a single best measure is impossible.

Rather than make such a selection, the best procedure in future applications of the model would be to use its structure to obtain information on the trade-offs between the measures of most importance. (The discussion in the next section on integration and travel is based on this type of analysis.) For

example, one could insert a constraint on total travel distance, and use travel time as the objective function. By making a half dozen runs, each with a different value for total distance allowed, one could determine how much of an improvement in travel time can be expected for each additional amount of distance allowed. This would allow the choice among the measures to be made on the basis of reasonably complete information, rather than prejudged before any real comparison is possible.

Integration

Two measures of integration were used. The first was based on the permitted variation in racial mix at each school, as described earlier. This measure was employed directly in the model, as a constraint on the solution. The second measure was an index of desegregation, calculated by multiplying, for each school, the percentage of the school's enrollment which is white by the percentage of the district's black population which is enrolled in this school; taking the sum of these products for all schools; and then dividing this by the percentage of students in the district who are white.[6] This index did not enter the model, but was calculated manually for each solution. Values for the two measures corresponded quite closely in the case study, so the remainder of this discussion will use only the index primarily because it is easier to interpret.

Exhibit 4 shows the four travel cost measures plotted against the level of integration as measured by the desegregation index. Each of the lines on the graph is drawn against a different scale on the vertical axis; the four appear on the same figure to facilitate comparisons. It is apparent that the four cost measures are strictly increasing functions of the level of desegregation. Further, the rate of increase in the function increases as the

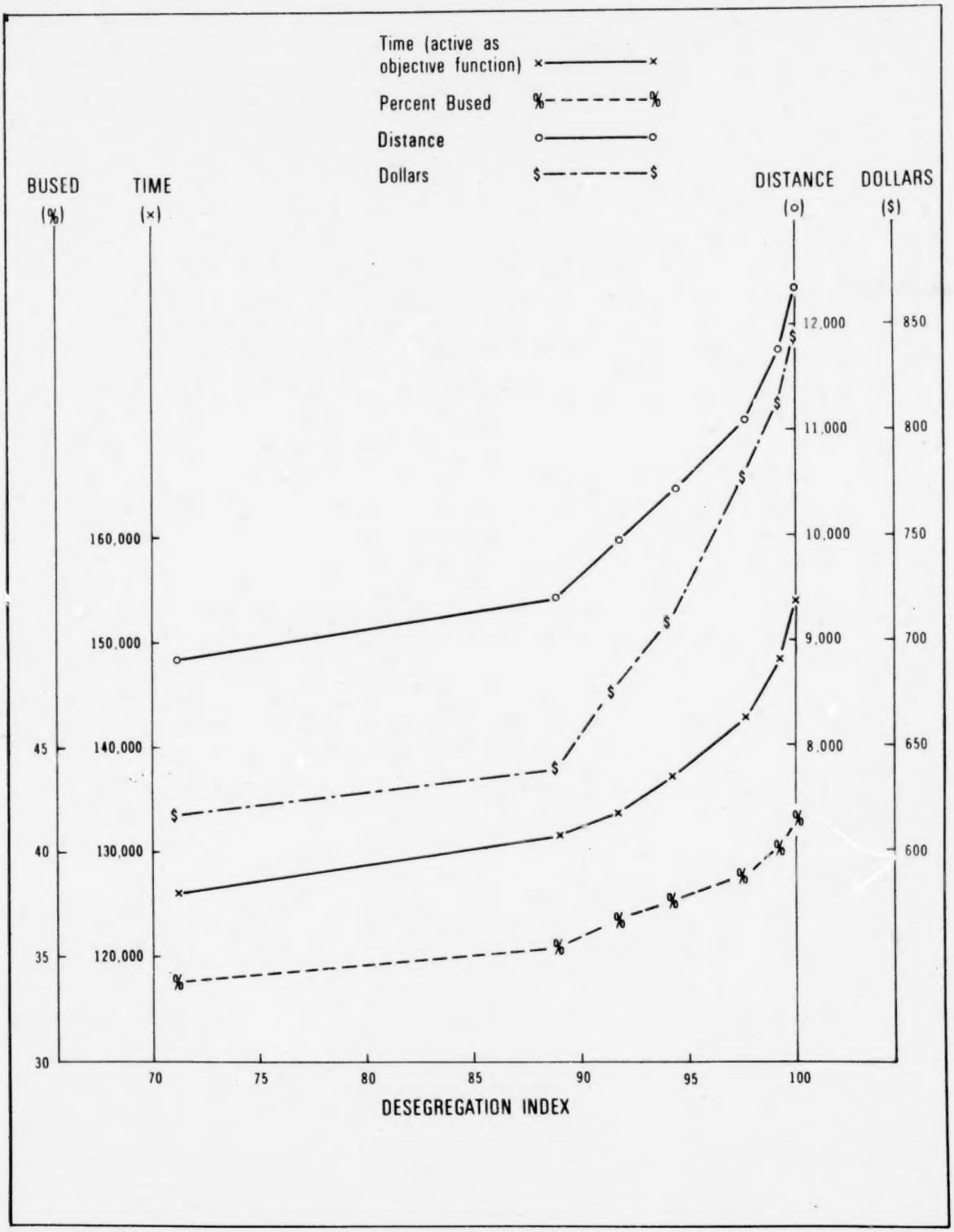


Exhibit 4. Travel costs as a function of the desegregation index.

value of the desegregation index gets larger. This means that each additional increment of integration achieved will cost more than the preceding increment. The other item of interest about these graphs is that there is not a strong breakpoint on most of them, at which costs suddenly rise much more steeply. There might be one at a value of roughly 88% on the distance and dollars curves, but the distance curve might also be interpreted as having its breakpoint at 97 or 98%. The curve for time is certainly a smooth curve, without a break--and this was the only cost measure which was actually minimized for this investigation. The other curves might be equally as smooth if they represented optimal solutions for those measures rather than simply the costs associated with the minimum time solution. The trade-off analysis, then, did not determine a best level of integration (as defined by a value at which costs suddenly increase rapidly), but it did provide useful information about the costs of achieving different levels of integration.

Limits on individual travel

Some of the objections to busing students to integrate schools have been based on the hardships this causes individuals, particularly in terms of the time used for travelling, or in terms of the distance to be overcome for parental conferences, for taking sick children home, and so forth. Two variations of the model investigated the effect which introducing upper limits on individual travel would have on the total travel costs, and on the spatial pattern of new schools and resulting allocations.

The costs of these limitations can be seen in Exhibit 5. A thirty minute limit on individual travel results in increases of just under 20% in both total travel time and percentage bused, and much smaller increases in total distance

(6%) and monetary cost (3 1/3%). Costs with a three mile limit on any individual's travel to school behave similarly, with 14 to 15% increases in total time and percentage bused and 2 to 3% increases in distance and monetary cost.

Exhibit 5

Total costs under limitations on individual travel

Cost measure	Maximum travel allowed		
	No limits	30 minutes	3 miles
Time	137,099	164,407	156,595
Distance	10,439	11,077	10,641
Dollars	709.0	732.7	728.3
Percentage bused	37.64	45.06	43.29

More important than these increases in total travel costs, however, are the differences in new school locations determined under these limitations. Both the thirty minute and three mile restrictions force one of the new schools to be located toward the north of the district, with a consequent change in the location of the second school as well. (Compare Exhibits 3 and 6.) As location choices are considerably more permanent than are attendance areas, maximum travel limits should be used only if it is certain that they are long-term criteria--that opinion about what constitutes a long journey to school will not change.

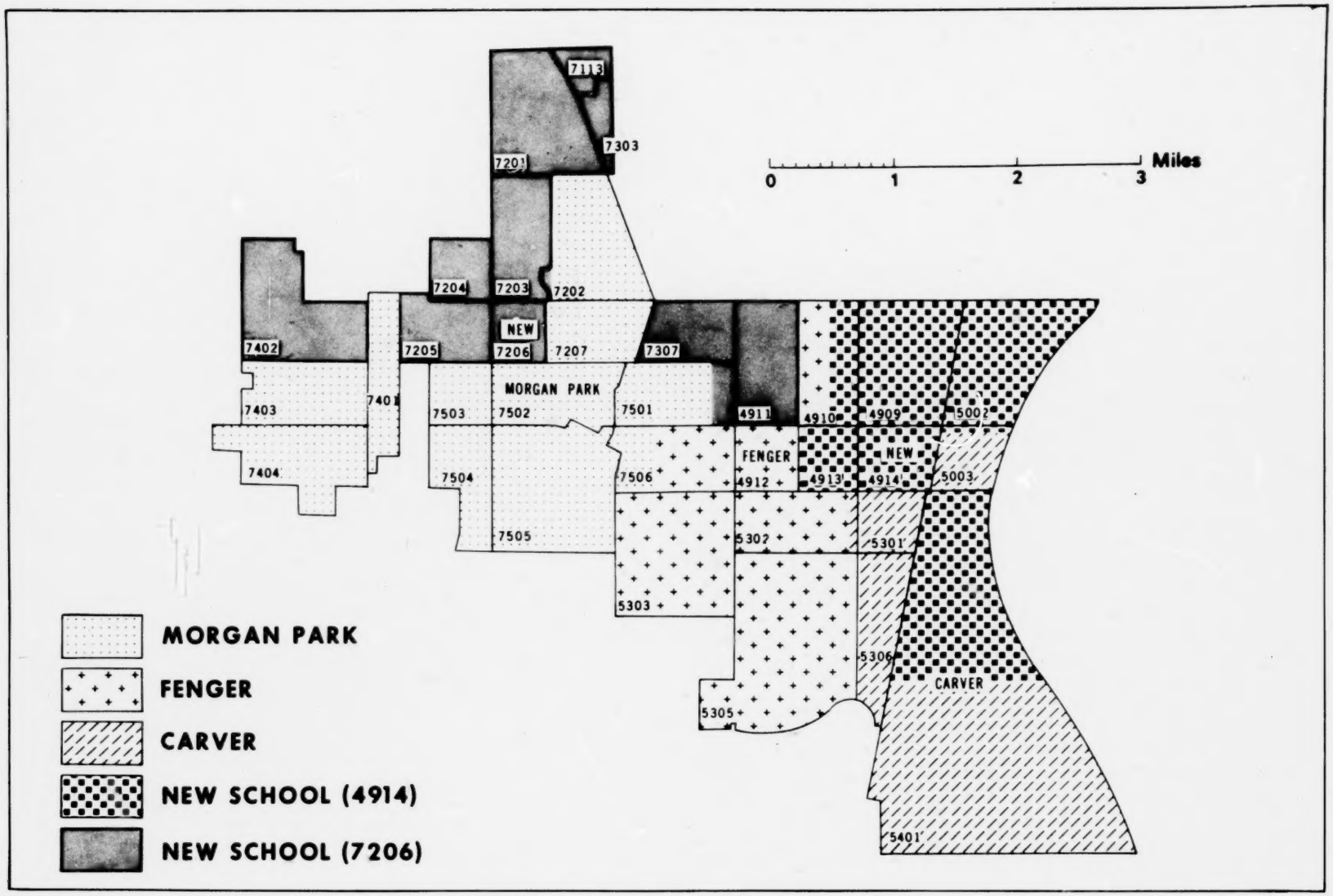


Exhibit 6. Optimal solution under a maximum travel limit of 30 minutes.

SUMMARY AND IMPLICATIONS

The work reported here had two main purposes. The first was to investigate possible criteria for locating new schools, and for allocating students to the full set of schools, both new and old. The second was to evaluate a particular model for determining optimal locations and allocations. The model was judged not only on its ability to provide information about the criteria being investigated, but also on its general applicability to and usefulness in school planning situations.

As developed, the location-allocation model is based on the branch of mathematical programming which deals with mixed integer problems. Linear programming techniques have been used for several years in school districting problems. The addition of the integer part of the problem permits solution for locations as well as for the allocations, or attendance areas. Programming models are formulated in terms of an objective function, which is to be either maximized or minimized, and a set of constraints, which contain limitations on the variables.

Three different location criteria were identified for analysis with this model: the total travel by students to school (measured in time, distance, dollars, or percentage who were bused to school), which served as the objective function to be minimized; racial integration of the system, which was employed as a constraint on the solution; and upper limits on the amount of travel an individual had to make to get to school (measured in time or distance), which also entered as a constraint. In addition to these constraints, there were two others imposed by the nature of the problem: each student must be assigned to one and only one school; and no school may be assigned more students than its [illegible] capacity. Eleven variations of the model were solved, to provide

sufficient data about the location criteria. A typical variation would be expressed as follows. Determine new locations for two schools, each of 1500 capacity, and allocations of students to all schools, such that the total time spent by all students travelling to school is minimized, subject to these constraints:

1. each student must be assigned to a single school;
2. no school can have more students than it has capacity;
3. the racial mix at each school shall not vary by more than 15% from the average in the area;
4. no student may travel more than 30 minutes to get to school.

The case study employed for testing the model was based on 1970 data for school district 18 in the southwestern part of the city of Chicago. This district had 7874 public high school students attending three high schools with a combined nominal capacity of 5070. With the exception of two atypical attributes, the district was a reasonably representative sample of the problems faced by school planners: the existence of residential segregation; a sprawling, non-compact area; large size; and a need for new schools. The atypical factors were the racial mix in the public high school population, which was 55% black and only 45% white, and the fact that school buses do not have to be provided. There is thorough coverage of the area by bus routes of the Chicago Transit Authority, and students are able to ride these for a reduced fare.

The most striking finding about the location criteria to come from this case study relates to the trade-offs between travel costs and integration. Each additional increment of integration increases travel costs more than does the previous increment, no matter which measures are used for travel cost and

for integration. The only other trade-off explicitly examined dealt with increases in total travel costs when upper limits were placed on individual travel. Here it was found that increases did occur, but that more important was the fact that new school locations in quite different parts of the district were selected. Because of the permanence of school locations, any use of such limitations on individual travel needs careful consideration. The four measures of total travel investigated were found to differ noticeably in their effects, although more with regard to allocations than to locations, but no single one of them stands out as any better than the others. It appears that the best procedure would be to use the model to identify trade-offs among them, as was done for integration and travel.

Evaluation of the model

Based on the findings from the case study, the model is definitely a useful addition to the list of planning techniques. Its principal use should probably be to provide information on the trade-offs among criteria relevant to a particular problem, as suggested for the several total travel measures, and done here for the integration and travel criteria. The reasoning behind this recommendation is that the technique, because it is rigorous and analytical, is rigid and uncompromising, and therefore should not be used to make an actual location selection. It can, however, provide quite useful information for that decision, which cannot be obtained easily any other way.

The question remains as to whether the model is too expensive to use. Experience in the case study indicates that it is not. Total computer costs for the eleven runs reported here were under five hundred dollars. This included data preparation, several false starts in which minor errors of formulation had to be corrected, and the final production runs. While solution procedures for mixed integer programming models are still in their early stages, and there

is no guarantee that every problem can be solved for similar costs, the programs do exist, and are operating in several places, so that this remains a reasonable estimate of the costs which might be incurred. (These costs can be placed in perspective if it is recalled that the optimal solution to the total travel costs resulted in one-way out of pocket expenditures for students of $455.60 per day.) As computation costs are directly related to both the total number of variables and the number of integer variables, restricting the choice of new locations to perhaps a half-dozen sites could reduce costs considerably. The case study described here permitted each of the 34 census tracts to be a potential new school location, which meant the problem contained 1122 regular variables plus 34 integer variables. If only 6 sites were to be considered, students from 100 different tracts (or smaller areal units) could be handled, and the problem would contain only 600 regular and 6 integer variables, which should result in a much less expensive analysis.

The question of the areal units to use in such a study is one of the problems which will be faced in any application of this model. For the case study, census tracts were used, because data were readily available in those units. However, geographically coded data on actual school populations would make a much better data base, and could help overcome the necessity for assumptions about uniformity of the population distribution within a census tract. Work on geo-coding is progressing in Chicago and other cities, and should help to make future applications of this model more reliable for planning.

Unfortunately, the existence of good geo-coding data will not totally surmount the problems of planning horizons, prediction, and planning over time mentioned earlier. That is, how many years ahead should school planning

be done--10, 20, 30, or what? Even if that can be decided, how accurate can population predictions be for small areas such as are needed for this model? And finally, even if we can obtain good population forecasts, how do we select locations which are best over time--not simply best at twenty years in the future, but best, on average perhaps, over the full twenty year span? But these are issues which the planner must face no matter which techniques he uses. They are not unique to this model, but are simply noticed more forcibly here because of the analytical nature of the model.

The final problem to be mentioned regarding use of the model is that each actual situation will contain its own characteristics, to which the model must be adapted. For example, in the case study, the capacity of Carver High School was not even sufficient to allow all of its own tract to be assigned there. This meant that the population of that tract had to be treated as if it came from two tracts--and the sizes of these varied when the permitted racial mix changed at Carver. However, the model is easily revised to encompass such problems, and the cost of doing so has been included in the estimates discussed above.

Implications for policy

The particular case study results (new school locations and attendance areas) are perhaps of limited value for planning new schools in Chicago, primarily because of the use of present population figures rather than projections for the future. They may, nevertheless, provide some insight into the problem as it now exists. More important are the case study findings regarding costs and cost trade-offs, because of the general forms of relationships which they indicate.

Even more important than these specific findings about the location criteria, both for Chicago and elsewhere, however, is the demonstration that this model works, for a reasonable cost, and can provide details about location considerations that could not be obtained any other way. Assuming that the kinds of criteria which were investigated in this case study are useful inputs to location decisions, the major implication of this report would seem to be that there is now a technique available for obtaining data on such criteria which should be used in preliminary planning studies.

NOTES

1. See, for example, the list on p. 6 of School Site Selection--A Guide, by R. C. Schneider, C. E. Wilsey, and SPL Staff (Stanford: Stanford University School of Education, 1961).

2. Council of Educational Facility Planners, Guide for Planning Educational Facilities (Columbus, O., 1971), pp. 58-59.

3. M. Yeates, "Hinterland delimitation: a distance minimizing approach," Professional Geographer, Vol. 15, No. 6 (1963), pp. 7-10.

4. Examples of this type of work are "School rezoning to achieve racial balance: a linear programming approach" by L. B. Heckman and H. M. Taylor, in Socio-Economic Planning Sciences Vol. 3 (1969), pp. 127-133; and "An operations research approach to racial desegregation of school systems" by S. Clarke and J. Surkis, in Socio-Economic Planning Sciences, Vol. 1 (1968), pp. 259-272.

5. Note that the racial categories shown were the only ones available in the preliminary census data, namely Negro and non-Negro.

6. This index has been taken from the Lambda Corporation report, School Desegregation with Minimum Busing (Arlington, Va., 1971), pp. 17-18.